



The Commonwealth of Massachusetts  
Executive Office of Health and Human Services  
Department of Public Health  
250 Washington Street, Boston, MA 02108-4619

DEVAL L. PATRICK  
GOVERNOR

TIMOTHY P. MURRAY  
LIEUTENANT GOVERNOR

JUDYANN BIGBY, MD  
SECRETARY

JOHN AUERBACH  
COMMISSIONER

## **Implementing New Race, Ethnicity, and Language Data Collection Standards**

### **Background and Rationale**

Addressing disparities in health and health care outcomes requires a deeper understanding of the populations we serve. Collecting detailed race, ethnicity, and language data are a means to providing this critical information. In 1997, the Federal Office of Management and Budget (OMB) promulgated revised standards for the collection of race and ethnicity data for use by Federal and state government agencies. These standards served as the basis for data collection in the 2000 Census and have been adopted by numerous public and private organizations for use in health care and other disciplines. MDPH has worked diligently to adapt these standards to the data collection needs of the Commonwealth by expanding response categories relevant to Massachusetts and adding questions to collect data on ethnicity and language.

In order to track and ultimately eliminate health disparities it is crucial to collect both broad race information and detailed ethnicity data. Race data are useful for monitoring discrimination, equality of opportunity and treatment, and indirectly, institutional racism. These broader categories yield more stable rates, are consistent with OMB Federal standard, and more routinely collected. While the OMB standard only calls for Hispanic ethnicity, detailed ethnicity data have enormous value as well, particularly tracking sub group differences masked by broad race categories. Ethnicity, also known as heritage or ancestry, is a more consistently understood concept, less likely to change over time, and less context dependent. It is more useful for program targeting and development than broad race categories. Ethnicity also may be a better measure of cultural practices, provide improved sensitivity to linguistic needs, and add insight into acculturation. Language information is the key to enhancing communication: it identifies needs for interpreter services and may help understand cultural practices related to health behavior.

## Principles

The MDPH standards for the collection of race and ethnicity data are based on four principles. We will briefly describe the importance of each of these principles:

1. Encourage patient self-report in the registration process

Comment: Self-report, where possible, is the mode of race and ethnicity ascertainment recommended by OMB. Studies have shown that self-identification provides the most consistent and valid information. Two alternative ascertainment methods: proxy (or informant) and observation, are less rigorous and more prone to misclassification error. Observation places an undue burden on the data collector, and both proxy and observation can be inconsistent across person, setting, and time. A key activity in the successful implementation of the MDPH standards is training intake personnel to gather race and ethnicity data consistently. MDPH, the City of Boston and Massachusetts General Hospital have developed training materials specifically for race data collection in health care settings in Massachusetts. Also, there are many regional and national training resources available.

2. Allow for the selection of multiple race categories

Comment: The OMB standards recommend that the format for individuals to report more than one race should be multiple responses to a single question rather than a catch all “multiracial” category. A “multiracial” category would be difficult to interpret and analyze. We strongly urge all MDPH Centers to explore data collection methods that will allow respondents to select more than one category. For those programs with limited data collection capacity, this standard could be implemented, for example, by the addition of a second variable, which, along with the first question would allow almost 95% of all patients with more than one race to identify themselves. For example, in Boston and Worcester, 95% of those who specified two or more races in Census 2000, had exactly two races.

3. Collect information on detailed ethnicity groups as well as broad race categories

*Race* is defined as the group or groups that you identify with as having similar physical characteristics or similar social and geographic origins. *Ethnicity* refers to your background, heritage, culture, ancestry, or sometimes the country where you or your family were born.

Comment: The Federal standards recognize only two ethnicities: Hispanic, and Non-Hispanic. The MDPH has found that there is a need for a greater number of ethnicities to be collected. Research at MDPH has found that there are significant differences in health outcomes *within* all broad race categories by ethnicity group. These differences are obscured when only broad race groups are examined. MDPH has also found that ethnicity data rather than race data are more useful for

program development, targeting interventions, expanding access to health services, and monitoring disparities. Massachusetts has found that different ethnicities have different health behaviors and health outcomes, thus we are leading the movement to collect detailed ethnicity data to better serve all residents.

4. Maintain consistency with Federal OMB standards

Comment: We recognize that it is important to collect data that comply with the 1997 OMB standards. Standardized race and ethnicity data will allow for direct comparisons within the state as well comparisons with other states and Federal databases.

Below is the paper version of the MDPH standard for race, ethnicity, and language data collection. This is to be considered the minimum set of categories for 'paper and form data collection. Computer data collection systems have been designed with a much more extensive list of ethnicities programmed into data dictionaries and look-up tables for ease of data entry. The goal of this effort remains collecting key data to monitor health disparities while minimizing the burden on data collectors, patients, and program participants.

Results from the pilot tests have been extremely encouraging. Training data collectors makes them much more comfortable asking these sensitive questions since they understand the importance of this information, how it will be used, how to respond to questions, and how to overcome uncertainty and resistance. We have also found that collecting detailed ethnicity and language preference allows program participants and patients to more fully reflect their heritage and provides more satisfaction than merely identifying broad race categories.

## MDPH Race-Ethnicity and Language Preference Data Collection Instrument: Option 1

**Introduction:** *In order to guarantee that all patients receive the highest quality of care and to ensure the best services possible, we are collecting data on race and ethnicity. Could you please select the category or categories that best describes your background?*

### 1. Are you Hispanic/Latino/Spanish?

- ☐ Yes
- ☐ No

### 2. What is your ethnicity? (You can specify one or more)

- |                                                             |                                                             |
|-------------------------------------------------------------|-------------------------------------------------------------|
| <input type="checkbox"/> African (specify _____)            | <input type="checkbox"/> Haitian                            |
| <input type="checkbox"/> African American                   | <input type="checkbox"/> Honduran                           |
| <input type="checkbox"/> American                           | <input type="checkbox"/> Japanese                           |
| <input type="checkbox"/> Asian Indian                       | <input type="checkbox"/> Korean                             |
| <input type="checkbox"/> Brazilian                          | <input type="checkbox"/> Laotian                            |
| <input type="checkbox"/> Cambodian                          | <input type="checkbox"/> Mexican, Mexican American, Chicano |
| <input type="checkbox"/> Cape Verdean                       | <input type="checkbox"/> Middle Eastern (specify _____)     |
| <input type="checkbox"/> Caribbean Islander (specify _____) | <input type="checkbox"/> Portuguese                         |
| <input type="checkbox"/> Chinese                            | <input type="checkbox"/> Puerto Rican                       |
| <input type="checkbox"/> Colombian                          | <input type="checkbox"/> Russian                            |
| <input type="checkbox"/> Cuban                              | <input type="checkbox"/> Salvadoran                         |
| <input type="checkbox"/> Dominican                          | <input type="checkbox"/> Vietnamese                         |
| <input type="checkbox"/> European                           | <input type="checkbox"/> Other (specify _____)              |
| <input type="checkbox"/> Filipino                           | <input type="checkbox"/> Unknown/not specified              |
| <input type="checkbox"/> Guatemalan                         |                                                             |

### 3. What is your race? (You can specify one or more)

- ☐ American Indian/Alaska Native (specify tribal nation \_\_\_\_\_)
- ☐ Asian
- ☐ Black
- ☐ Native Hawaiian or other Pacific Islander (specify \_\_\_\_\_)
- ☐ White
- ☐ Other (specify \_\_\_\_\_)
- ☐ Unknown/not specified

### 4. What language do you prefer to speak with us about health?

- |                                              |                                                          |
|----------------------------------------------|----------------------------------------------------------|
| <input type="checkbox"/> English             | <input type="checkbox"/> Somali                          |
| <input type="checkbox"/> Spanish             | <input type="checkbox"/> Arabic                          |
| <input type="checkbox"/> Portuguese          | <input type="checkbox"/> Albanian                        |
| <input type="checkbox"/> Cape Verdean Creole | <input type="checkbox"/> Chinese (specify dialect _____) |
| <input type="checkbox"/> Haitian Creole      | <input type="checkbox"/> Russian                         |
| <input type="checkbox"/> Khmer               | <input type="checkbox"/> Other (specify _____)           |
| <input type="checkbox"/> Vietnamese          |                                                          |

### 5. What language do you prefer to read health-related materials? \_\_\_\_\_

## MDPH Race-Ethnicity and Language Preference Data Collection Instrument: Option 2

**Introduction:** *In order to guarantee that all patients receive the highest quality of care and to ensure the best services possible, we are collecting data on race and ethnicity. Could you please select the category or categories that best describes your background?*

### 1. What is your ethnicity? (You can specify one or more)

- |                                                             |                                                             |
|-------------------------------------------------------------|-------------------------------------------------------------|
| <input type="checkbox"/> African (specify _____)            | <input type="checkbox"/> Haitian                            |
| <input type="checkbox"/> African American                   | <input type="checkbox"/> Honduran                           |
| <input type="checkbox"/> American                           | <input type="checkbox"/> Japanese                           |
| <input type="checkbox"/> Asian Indian                       | <input type="checkbox"/> Korean                             |
| <input type="checkbox"/> Brazilian                          | <input type="checkbox"/> Laotian                            |
| <input type="checkbox"/> Cambodian                          | <input type="checkbox"/> Mexican, Mexican American, Chicano |
| <input type="checkbox"/> Cape Verdean                       | <input type="checkbox"/> Middle Eastern (specify _____)     |
| <input type="checkbox"/> Caribbean Islander (specify _____) | <input type="checkbox"/> Portuguese                         |
| <input type="checkbox"/> Chinese                            | <input type="checkbox"/> Puerto Rican                       |
| <input type="checkbox"/> Colombian                          | <input type="checkbox"/> Russian                            |
| <input type="checkbox"/> Cuban                              | <input type="checkbox"/> Salvadoran                         |
| <input type="checkbox"/> Dominican                          | <input type="checkbox"/> Vietnamese                         |
| <input type="checkbox"/> European                           | <input type="checkbox"/> Other (specify _____)              |
| <input type="checkbox"/> Filipino                           | <input type="checkbox"/> Unknown/not specified              |
| <input type="checkbox"/> Guatemalan                         |                                                             |

### 2. What is your race? (You can specify one or more)

- ☐ American Indian/Alaska Native (specify tribal nation \_\_\_\_\_)
- ☐ Asian
- ☐ Black
- ☐ Hispanic/Latino/Black
- ☐ Hispanic/Latino/White
- ☐ Hispanic/Latino/other
- ☐ Native Hawaiian or other Pacific Islander (specify \_\_\_\_\_)
- ☐ White
- ☐ Other (specify \_\_\_\_\_)
- ☐ Unknown/not specified

### 3. What language do you prefer to speak with us about health?

- |                                              |                                                          |
|----------------------------------------------|----------------------------------------------------------|
| <input type="checkbox"/> English             | <input type="checkbox"/> Somali                          |
| <input type="checkbox"/> Spanish             | <input type="checkbox"/> Arabic                          |
| <input type="checkbox"/> Portuguese          | <input type="checkbox"/> Albanian                        |
| <input type="checkbox"/> Cape Verdean Creole | <input type="checkbox"/> Chinese (specify dialect _____) |
| <input type="checkbox"/> Haitian Creole      | <input type="checkbox"/> Russian                         |
| <input type="checkbox"/> Khmer               | <input type="checkbox"/> Other (specify _____)           |
| <input type="checkbox"/> Vietnamese          |                                                          |

### 4. What language do you prefer to read health-related materials? \_\_\_\_\_

# MDPH Detailed Ethnicity Categories and Supplemental Code Set

December 2006

| ETHNICITY CATEGORIES                      | Subcategories and Supplemental Code Set                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Cuban 2182-4                              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Dominican 2184-0                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Mexican, Mexican American, Chicano 2148-5 | Mexican American 2149-3, Mexicano 2150-1, Chicano 2151 – 9 , La Raza 2152-7, Mexican American Indian 2153-5                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Puerto Rican 2180-8                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Salvadoran 2161-8                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Central American (Other) 2155-0           | Costa Rican 2156-8, Nicaraguan 2159-2, Panamanian 2160-0, Central American Indian 2162-6, Belize                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| South American (Other) 2165-9             | Argentinean 2166-7, Bolivian 2167-5, Chilean 2168-3<br>Ecuadorian 2170-9, Paraguayan 2171-7, Peruvian 2172-5, Uruguayan 2173-3, Venezuelan 2174-1, South American Indian 2175-8, Criollo 2176-6, Guyana                                                                                                                                                                                                                                                                                                                                                                                                                             |
| African 2060-2                            | Botswanan 2061-0, Ethiopian 2062-8, Liberia 2063-6, Namibian 2064-4, Nigerian 2065-1, Zairean 2066-9<br>African also includes: Angola, Benin, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Comoros, Congo, Cote d'Ivoire, Djibouti, Egypt, Equatorial Guinea, Eritrea, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Libya, Madagascar, Malawi, Mali, Mauritania, Mauritius, Morocco, Mozambique, Niger, Reunion, Rwanda, Sao Tome & Principe, Senegal, Seychelles, Sierra Leone, Somalia, South Africa, Sudan, Swaziland, Tanzania, Togo, Tunisia, Uganda, Western Sahara, Zambia, and Zimbabwe |
| African American 2058-6                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| American AMERCN                           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Asian 2028-9                              | Bangladeshi 2030-5, Bhutanese 2031-3, Burmese 2032-1, Hmong 2037-0, Iwo Jiman 2048-7, Indonesian 2038-8<br>Madagascar 2052-9, Malaysian 2042-0, Maldivian 2049-5, Nepalese 2050-3, Okinawan 2043-8, Pakistani 2044-6, Singaporean 2051-1, Sri Lankan 2045-3, Taiwanese 2035-4, Thai 2046-1                                                                                                                                                                                                                                                                                                                                          |
| Asian Indian 2029-7                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Brazilian BRAZIL                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Cambodian 2033-9                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Cape Verdean CVERDN                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Caribbean Island CARIB                    | Barbadian 2068-5, Dominica Islander 2070-1, Jamaican 2072-7, Trinidadian 2074-3, Tobagoan 2073-5, West Indian 2075-0                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| Chinese 2034-7                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Columbian 2169-1                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| European 2108-9                           | English 2110-5, French 2111-3, German 2112-1, Irish 2113-9, Italian 2114-7, Scottish 2116-2, Greek GRK, Spanish SPAN, Armenian 2109-7, Polish 2115-4<br>Albanian ALBA, Azerbaijan AZER, Belarus BELA, Bosnia and Herzegovina BOSHER, Bulgaria BULG, Croatia CRO, Czech Republic CZECH, Estonia EST, Georgia GEOR, Hungary HUNG, Latvia LAT, Lithuania LITH, Moldova MOLD,                                                                                                                                                                                                                                                           |

|                                        |                                                                                                                                                                                                                                                                                      |
|----------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                        | Macedonia MACD, Montenegro MONT, Romania ROM, Serbia SERB, Slovakia SLOVK, Slovenia SLOVE, and Ukraine UKR                                                                                                                                                                           |
| Filipino 2036-2                        |                                                                                                                                                                                                                                                                                      |
| Guatemalan 2157-6                      |                                                                                                                                                                                                                                                                                      |
| Haitian 2071-9                         |                                                                                                                                                                                                                                                                                      |
| Honduran 2158-4                        |                                                                                                                                                                                                                                                                                      |
| Japanese 2039-6                        |                                                                                                                                                                                                                                                                                      |
| Korean 2040-4                          |                                                                                                                                                                                                                                                                                      |
| Laotian 2041-2                         |                                                                                                                                                                                                                                                                                      |
| Middle Eastern or North African 2118-8 | Assyian 2119-6, Egyptian 2120-4, Iranian 2121-2, Iraqi 2122-0, Lebanese 2123-8, Palestinian 2124-6, Syrian 2125-3, Afghanistani 2126-1, Israeli 2127-9, Middle Eastern also includes: Algerian, Jordan, Kuwait, Oman, Qatar, Saudi Arabia, Sudanese, United Arab Emirates, and Yemen |
| Portuguese PORTUG                      | Azorean, Canarian 2145-1                                                                                                                                                                                                                                                             |
| Russian RUSSIA                         |                                                                                                                                                                                                                                                                                      |
| Vietnamese 2047-9                      |                                                                                                                                                                                                                                                                                      |
| Other OTHER                            |                                                                                                                                                                                                                                                                                      |
| Unknown/not specified UNKNOW           |                                                                                                                                                                                                                                                                                      |